METEOROLOGICAL OFFICE
28 OCT 1986
LIBRARY

Long-Range Forecasting and Climate Research

Lagged-average forecast experiments with a 5-level general circulation model

by

J.M. Murphy

LRFC 8

March 1986

ORGS UKMO L

National Meteorological Library
FitzRoy Road, Exeter, Devon. EX1 3PB

FHB

KW

LONG-RANGE FORECASTING AND CLIMATE RESEARCH MEMORANDUM NO. 8
LRFC 8

LAGGED-AVERAGE FORECAST EXPERIMENTS WITH A 5-LEVEL GENERAL CIRCULATION MODEL

By J M Murphy

March 1986

(This paper has not been published. Permission to quote from it must be obtained from the Assistant Director of the above Meteorological Office Branch).

Abstract

The Meteorological Office 5-level model has been used to investigate the lagged-average forecast (LAF) method of producing ensemble forecasts. Four independent winter LAF experiments were run, each containing seven individual 50-day integrations initialised from successive analyses 24 hours apart.

From the model's intrinsic and practical error growth statistics optimal forecast weights were calculated to maximise the skill of the LAF ensemble-mean. There was no improvement in skill compared with the most recent individual forecast in the experiments, however, since the practical forecast errors in the individual integrations within a 5-level model LAF ensemble are highly correlated with each other.

The weighting calculations were used to show that, given a better forecast model, an improvement in medium-range skill can be achieved with this method.

1. INTRODUCTION

It is a well documented feature of numerical weather prediction (NWP) models that pairs of integrations, differing from each other initially by only a small amount, will grow apart with time until eventually they are as widely separated as random states chosen from a suitable climatological distribution (e.g. Leith, 1974; Lorenz, 1982). The problem of extended-range forecasting using such models therefore takes on a probabilistic nature, since a number of forecasts initialised from separate analyses, each consistent with a particular set of observations, and equally likely to represent the true initial atmospheric state, will each give a different prediction for the evolution of the real atmosphere.

In the model phase space (Epstein, 1969; Gleeson, 1970; Leith, 1974), we may represent the distribution of possible initial states by a continuous probability density function ψ (t=0), centred about the true state with a spread reflecting the size of observation errors. Running forecasts from an infinite number of initial states sampled from ψ (o) gives the time evolution ψ (t) of the probability distribution for the state of the real atmosphere, under the assumption of a perfect forecast model (i.e. that given the correct initial state, the model will produce a perfect forecast).

We may obtain a sampling approximation, ψ' (t), to ψ (t) by running a finite ensemble of forecasts from initial conditions chosen at random from ψ (o). Under the assumption that ψ (t) evolves as a normal distribution (Leith, 1974; Seidman, 1981), the best-estimate forecast is the centroid of ψ (t), which we approximate using the centroid of our ensemble. The larger the ensemble size, the better the estimate we obtain of the centroid of ψ. Various 'perfect model' studies have confirmed that increased skill may be achieved by ensemble-averaging (Seidman, 1981; Hoffman and Kalnay, 1983; Murphy, 1986).

One problem we face in practice, however, is that ψ (o) is unknown. We have available only a single observed state, plus an idea of the typical errors inherent in the observing system and analysis method. A common method of generating alternative possible initial states is to perturb the observed state in some way consistent with known errors. This procedure has the shortcoming that the resulting distribution of states has the observed, rather than the true, state as its centroid. Furthermore it is a complex matter to account properly in the perturbation procedure for variations in the size of measurement errors for different components of the observed state, and also for off-diagonal elements of the error covariance matrix specifying correlations between the errors of pairs of components.

In practice, simplified random perturbation (RP) procedures have been adopted, in which the variables at each grid-point are perturbed independently by an amount consistent with space-averaged analysis errors (Spar et al, 1978; Seidman, 1981; Shukla, 1981). A variant on this procedure is to create the perturbed states from linear combinations of the observed state with independent analyses (Murphy, 1986). In this paper, however, we shall consider an alternative approach, called the lagged-average forecast (LAF) method (see below).

Whichever method we use, it must be recognised that the theoretical potential for increased skill available through ensemble-averaging will be eroded, in practice, to a degree dependent on the extent of the model's deficiencies. In this study we use a hemispheric version of the Meteorological Office 5-level GCM. Mansfield (1986) has shown that the forecast skill of this model is rather low, and ensemble forecast experiments using the RP approach (Murphy, 1986), have shown that ensemble-averaging only gives an improvement in skill in those cases where the model shows skill beyond the normal range of deterministic predictability. Thus we are unlikely to obtain a large improvement in skill using the LAF method either. It is still of interest, however, to compare the two approaches, since in future experiments with more skilful models, the question of which method gives better results will become important, as the benefits of ensemble forecasting, demonstrated in the perfect model experiments, become more fully realised in practice.

2. LAGGED-AVERAGE FORECAST METHOD

This consists of using forecasts made from a number of analyses, each lagging the start of the forecast period by a different amount (Hoffman and Kalnay, 1983). We may view this procedure as using the dynamical model to extrapolate previous observations up to the analysis time to obtain an estimate of ψ (o), (Miyakoda and Talagrand, 1971) and then using the subsequent evolutions of the individual integrations to form the ensemble forecast. Figure 1 shows schematically the difference between the LAF and RP procedures. The LAF method has the immediate practical advantage that all the information it uses is generated routinely in the course of normal operational forecast procedures, thus making it easier to implement than the RP approach. The overriding criterion of judgement however is the question of which method gives a better approximation to ψ (o).

A positive feature of the LAF method is that it generates a prediction of ψ (o) quite naturally without requiring us to estimate the spread of ψ (o) due to observation errors. Provided systematic errors are absent the centroid of ψ' (o), our sampling approximation to ψ (o), is an unbiased estimate of the true state. If the prediction errors incurred in extrapolating the time-lagged observed states up to the analysis time are small, the spread of ψ' (o) will accurately reflect the uncertainty of the initial state. Thus the LAF method may be superior to the RP method if the skill of the model is sufficiently high, and the interval τ between successive analyses sufficiently short, to ensure that prediction errors remain small (~ the observation errors) for a time $(M-1)\tau$, where M is the membership of the ensemble. Also important is the extent to which the errors in successive observed states are correlated. If the correlations are high the amount of independent information available for the determination of ψ (o) is reduced and the sampling error associated with ψ' (o) is correspondingly larger.

Hoffman and Kalnay (1983) studied the relative merits of the two methods using a simple model with idealised observing errors, finding that the LAF method gave slightly better forecasts. However the factors discussed above will vary from one operational forecast system to the next and the judgement of which method is superior is liable to vary correspondingly.

To demonstrate the LAF method ensemble forecasts with M=7 have been carried out using τ = 24 hours. This interval, enforced through having only midnight analyses available, is longer than desirable, since prediction errors in the extrapolation of lagged analyses will then be rather large. Certainly for medium-range forecasting, up to a range of about 10 days, a smaller value of 6 or 12 hours would be required, although with regard to forecasting on a monthly time-scale, for which we are concerned with long-wave predictability, 24 hours might be more acceptable, since error growth in the early stages of numerical forecasts is dominated by the synoptic scales (Boer, 1984).

In a LAF ensemble the individual forecasts made from the more recent analyses should receive higher weighting in the formation of the ensemble-mean due to the additional prediction errors carried by forecasts from older analyses. The weights, which are functions of time, are determined by demanding that the climatic average skill of the weighted ensemble-mean forecast is a maximum. In theory we might include the a priori predictability information carried by the second moment of $\psi'(t)$ in the calculation, in which case the weights would vary from one ensemble forecast to another, but this has not been attempted in the present study. The determination of the weights is described in the following section.

3. OPTIMAL COMBINATION FOR LAGGED-AVERAGE FORECASTS

Consider a set of M individual forecasts initialised at successive instants each separated by an interval τ (Figure 1). We wish to calculate weights for these integrations in a linear combination such that the climatic average forecast skill of the weighted ensemble-mean is a maximum. If $u_k(t+(k-1)\tau)$ represents the anomaly value of some model variable in forecast k, where forecast time t is measured from the initialisation time of the most recent of the M forecasts (denoted by k=1), the weighted ensemble-mean anomaly u(t) is given by

$$\hat{u}(t) = \sum_{k=1}^{M} p_k(t)u_k(t+(k-1)\tau) / \sum_{k=1}^{M} p_k(t) \qquad (1)$$

where the weights $p_k(t)$ are normalised such that $\sum_{k=1}^{M} p_k(t)=1$. In general the weights will depend on which objective skill score is used to measure forecast predictability. We calculate weights corresponding to two measures of skill, namely normalised error variance and anomaly correlation.

Normalised error variance

We shall use the notation $\langle\ \rangle$ to denote an average over a large number of forecast experiments using initial conditions chosen at random from a suitable climate distribution.

The climatic average normalised error variance <E> of a weighted ensemble-mean forecast is given by

$$\langle E\rangle = \frac{\langle(\hat{u}-u_o)^2\rangle}{\langle(\hat{u})^2\rangle+\langle u_o^2\rangle} = 1 - 2\,\frac{\langle\hat{u}u_o\rangle}{\langle(\hat{u})^2\rangle+\langle u_o^2\rangle} \qquad (2)$$

where u_o denotes the true state anomaly.
For ease of notation explicit time dependences are omitted in (2) and following equations. By substituting equation (1) into (2) we may show that

$$\langle E\rangle = 1 - \frac{2\sum_{k=1}^{M} p_k\langle u_k u_o\rangle / \sum_{k=1}^{M} p_k}{(\sum_{k=1}^{M} p_k^2\langle v_k\rangle + 2\sum_{k=1}^{M}\sum_{l=1}^{M} p_k p_l\langle u_k u_l\rangle)/(\sum_{k=1}^{M} p_k^2)+\langle v_o\rangle} \qquad (3)$$

where $\langle v_o\rangle = \langle u_o^2\rangle$ and $\langle v_k\rangle = \langle u_k^2\rangle$. In general the variances $\langle v_o\rangle$ and $\langle v_k\rangle$, k=1 → M, are all different due to model climate drift. Equation (3) may be recast by defining

$$w_k^2 = p_k^2\,\langle v_k\rangle/\langle v_o\rangle \qquad (4)$$

such that

$$\langle E\rangle = 1 - \frac{2\sum_{k=1}^{M} w_k\langle c_{ko}\rangle/(\sum_{k=1}^{M} w_k(\langle v_o\rangle/\langle v_k\rangle)^{1/2})}{(\sum_{k=1}^{M} w_k^2 + 2\sum_{k=1}^{M}\sum_{l=k+1}^{M} w_k w_l\langle c_{kl}\rangle)/(\sum_{k=1}^{M} w_k(\langle v_o\rangle/\langle v_k\rangle)^{1/2})^2 + 1} \qquad (5)$$

where $\langle c_{kl}\rangle = \langle u_k\, u_l\rangle/(\langle v_k\rangle\,\langle v_l\rangle)^{1/2}$ is the anomaly correlation between forecasts k and l and $\langle c_{ko}\rangle$ the anomaly correlation between forecast k and reality.

The weights w_k which minimise <E> may be found using the method of Lagrange multipliers, giving the M conditions

$$\frac{\partial(\langle E\rangle)}{\partial w_{k'}} + \lambda(\langle v_o\rangle/\langle v_{k'}\rangle)^{1/2} = 0, \text{ for } k' = 1 \to M,$$

where λ is the multiplier, plus the constraint $\sum_{k=1}^{M} w_k(\langle v_o\rangle/\langle v_k\rangle)^{1/2} = 1$.

Using equation (4) we recover the following M non-linear simultaneous equations for the M weights w_k,

$$(\langle c_{10}\rangle-\langle c_{ko}\rangle(\langle v_{k'}\rangle/\langle v_1\rangle)^{1/2})\,(1 + \sum_{k=1}^{M} w_k^2 + 2\sum_{k=1}^{M}\sum_{l=k+1}^{M} w_k w_l\langle c_{kl}\rangle)$$

$$= 2\,(\sum_{k=1}^{M} w_k\langle c_{ko}\rangle)(\sum_{k=1}^{M} w_k(\langle c_{k1}\rangle-\langle c_{kk'}\rangle(\langle v_{k'}\rangle/\langle v_1\rangle)^{1/2}), \qquad (6)$$

$$\text{for } k' = 2 \to M, \text{ and } \sum_{k=1}^{M} w_k(\langle v_o\rangle/\langle v_k\rangle)^{1/2} = 1$$

Thus in order to calculate the w_k, and hence the p_k, at some time level, we require the anomaly correlations $\langle c_{k0}\rangle$ between the individual forecasts and reality, the correlations $\langle c_{kl}\rangle$ between pairs of forecasts and the variance ratios $\langle v_k\rangle/\langle v_o\rangle$ between the model forecasts and reality.

Anomaly correlation

The climatic average anomaly correlation $\langle C\rangle$ of a weighted ensemble-mean forecast is given by

$$\langle C\rangle = \langle \hat{u}u_o/(\hat{u})^2 u_o^2)^{1/2}\rangle = \langle \hat{u}u_o\rangle/(\langle(\hat{u})^2\rangle\langle u_o^2\rangle)^{1/2}$$

assuming that climatic variations are negligible. Since no attempt is made to account for case-by-case variations in predictability, which would require the use of variable weights, this is a consistent assumption to make. From equations (1) and (4)

$$\langle C\rangle = \sum_{k=1}^{M} w_k\langle c_{ko}\rangle/\left(\sum_{k=1}^{M} w_k^2 + 2\sum_{k=1}^{M}\sum_{l=k+1}^{M} w_k w_l\langle c_{kl}\rangle\right)^{1/2} \qquad (7)$$

with the w_k defined as before. If we apply the method of Lagrange multipliers to equation (7) to find the values of w_k which maximise $\langle C\rangle$, we obtain the following conditions,

$$(\langle c_{1o}\rangle - \langle c_{k'o}\rangle)(\langle v_{k'}\rangle/\langle v_1\rangle)^{1/2})\left(\sum_{k=1}^{M} w_k^2 + 2\sum_{k=1}^{M}\sum_{l=k+1}^{M} w_k w_l\langle c_{kl}\rangle\right)$$

$$= \left(\sum_{k=1}^{M} w_k\langle c_{ko}\rangle\right)\left(\sum_{k=1}^{M} w_k(\langle c_{k1}\rangle - \langle c_{kk'}\rangle(\langle v_{k'}\rangle/\langle v_1\rangle)^{1/2}\right), \qquad (8)$$

$$\text{for } k' = 2 \to M, \text{ and } \sum_{k=1}^{M} w_k(\langle v_o\rangle/\langle v_k\rangle)^{1/2} = 1$$

Since these conditions are not identical to equations (6), the weights are different for the two skill scores.

Determination of weights

In general the error and variance statistics appearing in equations (6) and (8) are different for each model variable, so strictly the weights for each variable should be determined separately. In practice we have assumed the error and relative variance statistics to be identical for all variables and have used grid point-averaged values determined from daily model 500 mb height forecast fields to determine the weights.

The variance ratios $\langle v_o \rangle / \langle v_k \rangle$ were estimated from the model climatology (see below), using the day zero variance to estimate $\langle v_o \rangle$. The required practical forecast skill scores $\langle c_{ko} \rangle$ for each time level t were found from a mean skill curve obtained by verifying all seven individual forecasts in each of the four LAF ensembles against the real atmosphere. To find the "perfect model" scores $\langle c_{kl} \rangle$ an additional skill curve showing intrinsic error growth was obtained by repeating the above verification using additional integrations, created by perturbing the observed state for December 29 in the manner described in Murphy (1986), to represent nature.

By assuming (following Lorenz, 1982), that the intrinsic error growth rate depends only on the size of the error itself, the values of $\langle c_{kl} \rangle$ at a given time level t were deduced by reading off the values of $\langle c_{kl} \rangle$ at t=0 from the practical skill curve and extrapolating to time t using the perfect model skill curve.

For each time level equations (6) and (8) were solved numerically to determine the w_k values for normalised error variance and anomaly correlation respectively, from which corresponding weights p_k were found using equations (4).

Substituting these weights back into equations (5) and (7) enables us to find the expected improvement in the skill of the weighted ensemble-mean compared with that of the most recent individual forecast. This turns out to be very small (1-2%), for both normalised error variance and anomaly correlation.

Figures 2 and 3 show the time variation of the weights for the two skill scores. In both cases the weights do not become equal as $t \to \infty$. The reason is that whilst the individual skill curves $\langle c_{10} \rangle$, $\langle c_{20} \rangle$, ..., $\langle c_{M0} \rangle$ tend to zero at large forecast times, the ratios between them, such as $\langle c_{10} \rangle / \langle c_{20} \rangle$, do not. (This can be seen by considering equations (6) and (8) in the simple case of M=2).

4. NUMERICAL EXPERIMENTS

The model used was a hemispheric version of the UK Meteorological Office (UKMO) 5-level GCM (Corby et al, 1977) which has a quasi-uniform horizontal grid with a resolution of approximately 330 km. The initial data for the experiments was obtained from UKMO operational analyses. Further details of the model and initialisation procedure are given by Mansfield (1986). Climatological boundary forcing conditions were employed in all integrations.

A set of four wintertime LAF ensemble experiments, each containing seven individual integrations, was produced using initial data from 00Z December 23 to 00Z December 29 inclusive, in the years 1978-81. All individual integrations were run for 50 days.

In general the climate of a GCM drifts with time away from the true atmospheric climate as systematic errors in the model circulation develop. Mansfield (1986) has shown that this is certainly true of the 5-level model. In order to remove such systematic errors when verifying model forecasts against the real atmosphere an estimate of the model climatology, as a function of forecast time, is required. The integrations can then be

formulated as anomaly forecasts which are unbiased predictions of the evolution of the true atmospheric anomalies. A set of eight integrations, using initial data taken from seven different winters (see Murphy, 1986), was used for this purpose.

5. RESULTS

The calculation of forecast weights in section 3 reveals that the 5-level model's intrinsic and practical error growth statistics are such that we cannot expect the weighted ensemble-mean forecast to offer any significant increase in practical skill compared with the most recent individual forecast. In fact, from equations (5)-(8), it may be shown that no increase could be expected, even with a smaller value of τ or a larger ensemble size. The reason is that the rate at which the integrations within an ensemble diverge from each other is much smaller than the rate at which the integrations diverge from reality (see Murphy, 1986). In other words, the practical forecast errors in each successive integration are highly correlated with each other, so that ensemble-averaging does little to reduce them. Figure 4, which shows the skill of the daily weighted ensemble-mean forecast compared with that of the most recent individual forecast for the 500 mb height field in the area 30-85°N, averaged over the four LAF cases for days 1-14, confirms this experimentally, tying in also with results obtained from RP ensemble forecasts with the same model (Murphy, 1986).

However the skill of the weighted ensemble-mean exceeds that of the unweighted ensemble-mean, illustrating the need for a weighting procedure for forecast times up to the normal limit of deterministic predictability. Beyond this point there is no a priori reason to prefer one integration to another, so all should be given equal weight. (In fact none of the four LAF experiments show any skill beyond 14 days anyway).

Despite these results, it is possible that the LAF method may lead to increased medium-range predictability given a more skilful model. To demonstrate this the intrinsic and practical winter 500 mb height forecast error statistics given by Lorenz (1982), for the global, modified[1*] ECMWF model, were used to calculate the expected improvement in anomaly correlation available from an M=7 weighted ensemble-mean forecast, with τ = 12 or 24 hours (Figure 5). The results show that an increase in skill is obtained, particularly between days 6 and 10, during which period the anomaly correlation of the most recent individual forecast drops from 0.5 to 0.2.

If, following Lorenz, we adopt the point at which the rms forecast error reaches six sevenths of its limiting value as a measure of the deterministic predictability limit, the corresponding anomaly correlation limit of 0.27 is reached after 10.0 days with the weighted ensemble-mean forecast, compared to 8.6 days for the individual forecast, an increase of 15%. This result is liable to be somewhat optimistic, since in any particular case the forecast skill of the integrations, and the growth of

---

1
* The temporal mean and standard deviation of each spherical harmonic coefficient of the forecast field were adjusted to match the real climatology.

differences between them, will in general differ from the climatic average values on which the weights are based. If it turns out that there is some correlation between ensemble spread and practical forecast skill, as there is under perfect model conditions (Murphy, 1986), the weights could be varied case-by-case based on the spread, to account, as far as possible, for variations in practical predictability.

6. SUMMARY

A set of four 50-day lagged-average forecast (LAF) ensemble experiments, each containing seven individual integrations with an interval of 24 hours between successive analyses, was produced using a hemispheric version of the UK Meteorological Office 5-level GCM with climatological boundary forcing conditions. No skill was apparent in any of the experiments beyond a range of two weeks, taken to represent the normal limit of deterministic predictability.

For earlier forecast times it is necessary to weight each individual forecast in an LAF ensemble differently when forming the ensemble-mean, to take account of error growth occurring whilst the lagged analyses are extrapolated up to the start of the forecast period.

The weights, which are functions of forecast time, were calculated from the model's intrinsic and practical error growth statistics by demanding that average skill of the weighted ensemble-mean be maximised. Two sets of weights, corresponding to normalised error variance and anomaly correlation skill scores, were obtained. The calculations were used to confirm that no significant improvement in skill could be expected through ensemble-averaging, even if a shorter time-lagging interval or larger ensemble size were to be used. This is because, with the 5-level model, the practical forecast errors in each successive integration are highly correlated, so that little reduction in error can be achieved by ensemble-averaging. However, from the forecast error statistics for the ECMWF model given by Lorenz (1982), it appears that weighted LAF ensemble-mean forecasts can in theory yield some improvement in medium-range skill relative to individual forecasts, using more skilful NWP models.

References

Boer, G.J. 1984 A spectral analysis of predictability and error in an operational forecast system. Mon. Wea. Rev., 112, 1183-1197.

Corby, G.A., Gilchrist, A. and Rowntree, P.R. 1977 The United Kingdom Met Office 5-level general circulation model. Methods Comp Physics, 17, 67-110.

Epstein, E.S. 1969 Stochastic dynamic prediction. Tellus, 21, 739-759.

Gleeson, T.A. 1970 Statistical-dynamical predictions. J. Appl. Meteor., 9, 333-344.

Hoffman, R.N., and Kalnay, E.	1983	Lagged-average forecasting, an alternative to Monte Carlo forecasting. Tellus, 35, 100-118.
Leith, C.E.	1974	Theoretical skill of Monte Carlo forecasts. Mon. Wea. Rev., 102, 409-418.
Lorenz, E.N.	1982	Atmospheric predictability experiments with a large numerical model. Tellus, 34, 505-513.
Mansfield, D.A.	1986	The skill of dynamical long-range forecasts, including the effect of sea surface temperature anomalies. To appear in Q.J.R.M.S.
Miyakoda, K. and Talagrand, O.	1971	The assimilation of past data in dynamical analysis. I. Tellus, 23, 310-317.
Murphy, J.M.	1986	The impact of ensemble forecasts on predictability. In preparation.
Seidman, A.N.	1981	Averaging techniques in long-range weather forecasting. Mon. Wea. Rev., 109, 1367-1379.
Shukla, J.	1981	Dynamical predictability of monthly means. J. Atmos. sci., 38, 2547-2572.
Spar, J., Notario, J.J., and Quirk, W.J.	1978	An initial state perturbation experiment with the GISS model. Mon. Wea. Rev., 106, 89-110.

Figure Captions

Figure 1. Initial conditions for seven-member LAF and RP ensembles. The x-axis is forecast time and the y-axis schematically represents the model phase space. [——] denotes an observed state and [- - -] a state created by perturbing an observed state. τ is the interval between successive observed states.

Figure 2. Optimal forecast weights to minimise normalised error variance score of 5-level model M=7 LAF ensemble-mean, with time-lagging interval of 24 hours. The most recent individual forecast is denoted by 1.

Figure 3. Optimal forecast weights to maximise anomaly correlation score of 5-level model M=7 LAF ensemble-mean, with time-lagging interval of 24 hours. The most recent individual forecast is denoted by 1.

Figure 4. Daily forecast skill for 500 mb height field in the region 30-55°N, averaged over four LAF experiments. (——) skill of the most recent individual forecast (- - - -) skill of the weighted LAF ensemble-mean. (-x-x-) skill of the unweighted LAF ensemble-mean.

Figure 5. Theoretical estimate of the average improvement in the daily anomaly correlation score of the ECMWF model available from M=7 LAF ensemble-mean forecasts, deduced from the global, wintertime 500 mb height forecast error statistics of Lorenz (1982). (——) skill of the most recent individual forecast. (- - - -) skill of the weighted LAF ensemble-mean with time-lagging interval of 12 hours (-x-x-) skill of the weighted LAF ensemble-mean with time-lagging interval of 24 hours.

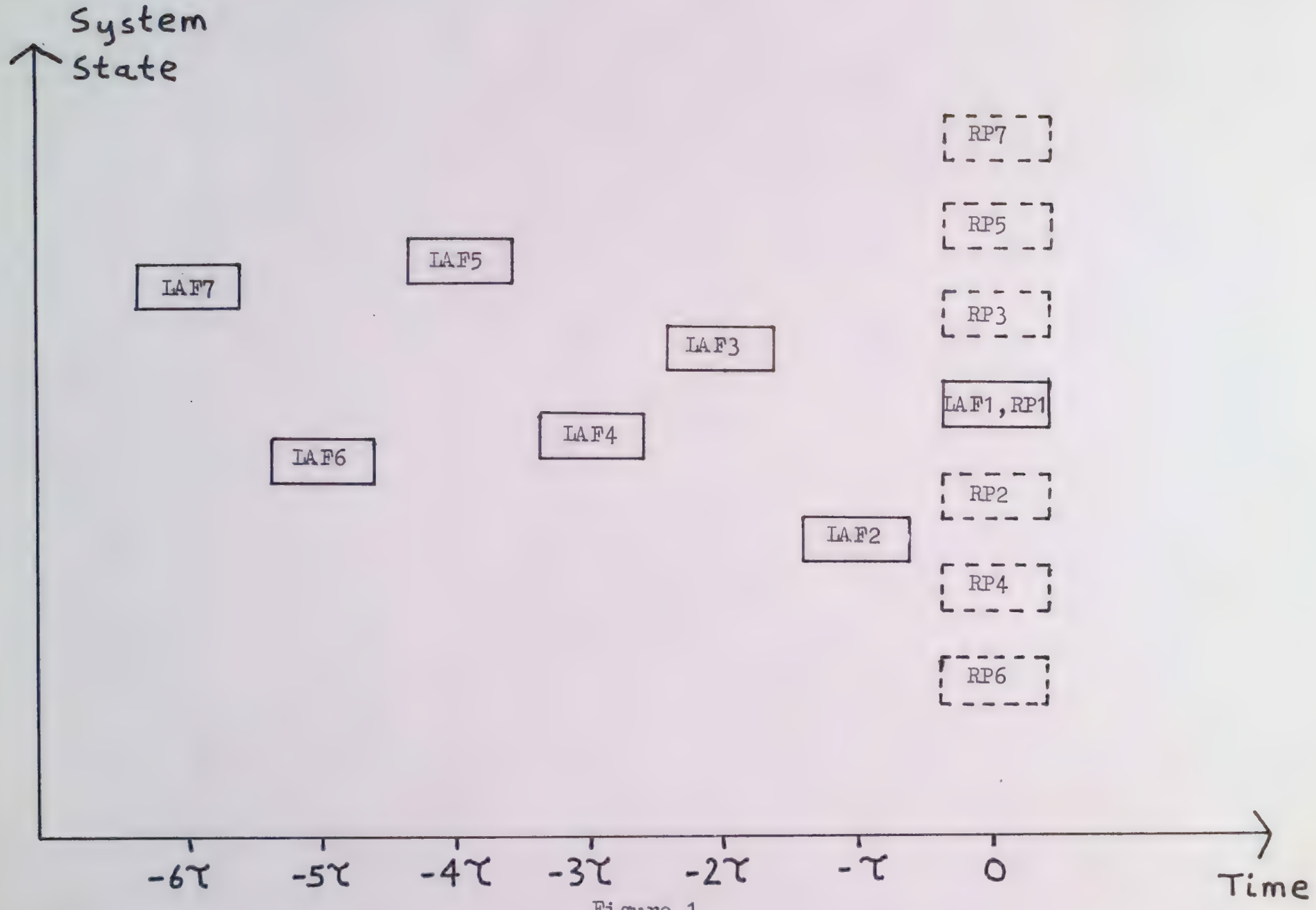

System State
RP7
RP5
RP3
LAF7
LAF5
LAF3
LAF1, RP1
LAF4
LAF6
RP2
LAF2
RP4
RP6
-6τ
-5τ
-4τ
-3τ
-2τ
-τ
0
Time

Figure 1

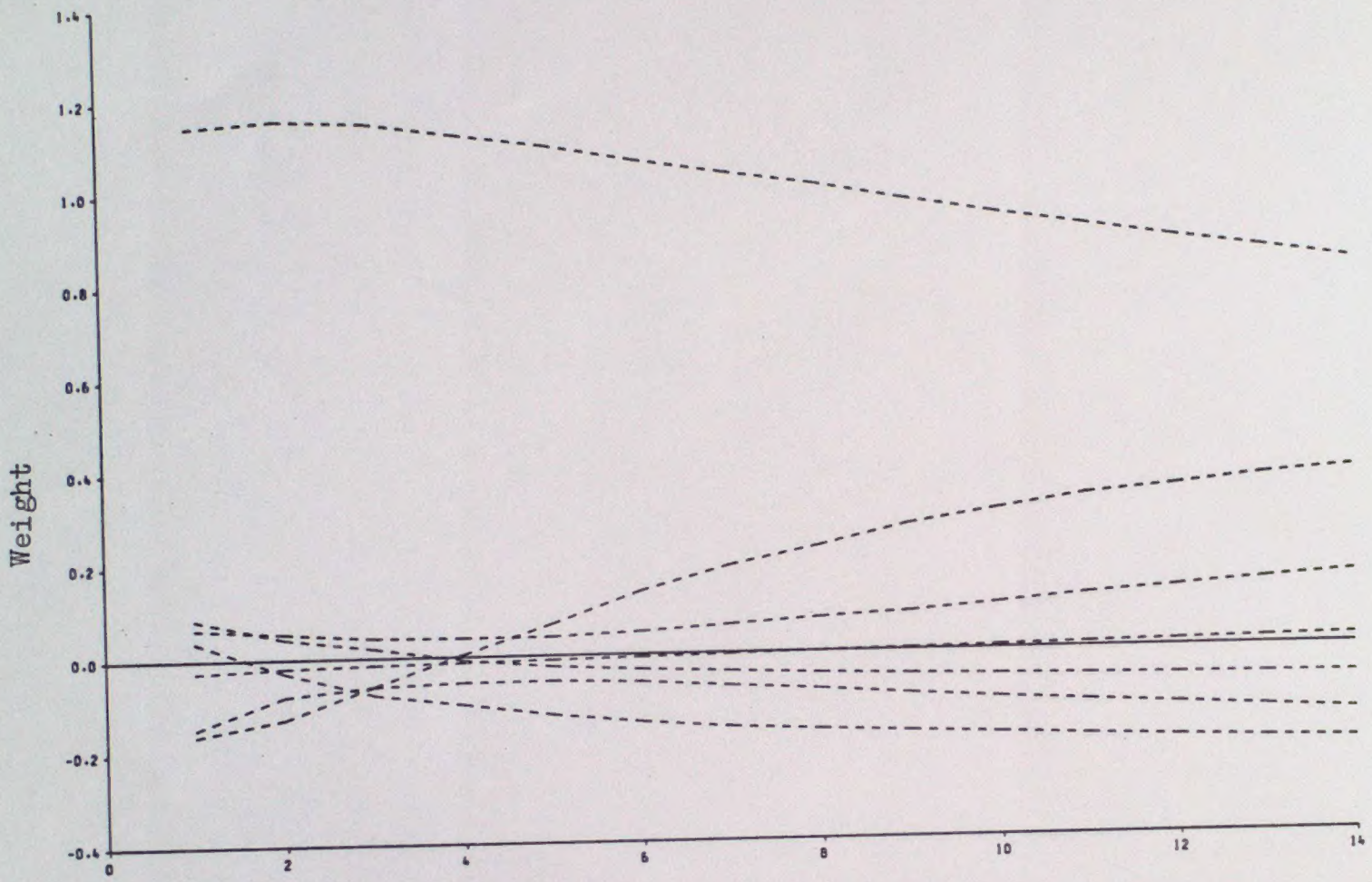

Weight
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
-0.2
-0.4
0
2
4
6
8
10
12
14
Forecast time (days)

Figure 2

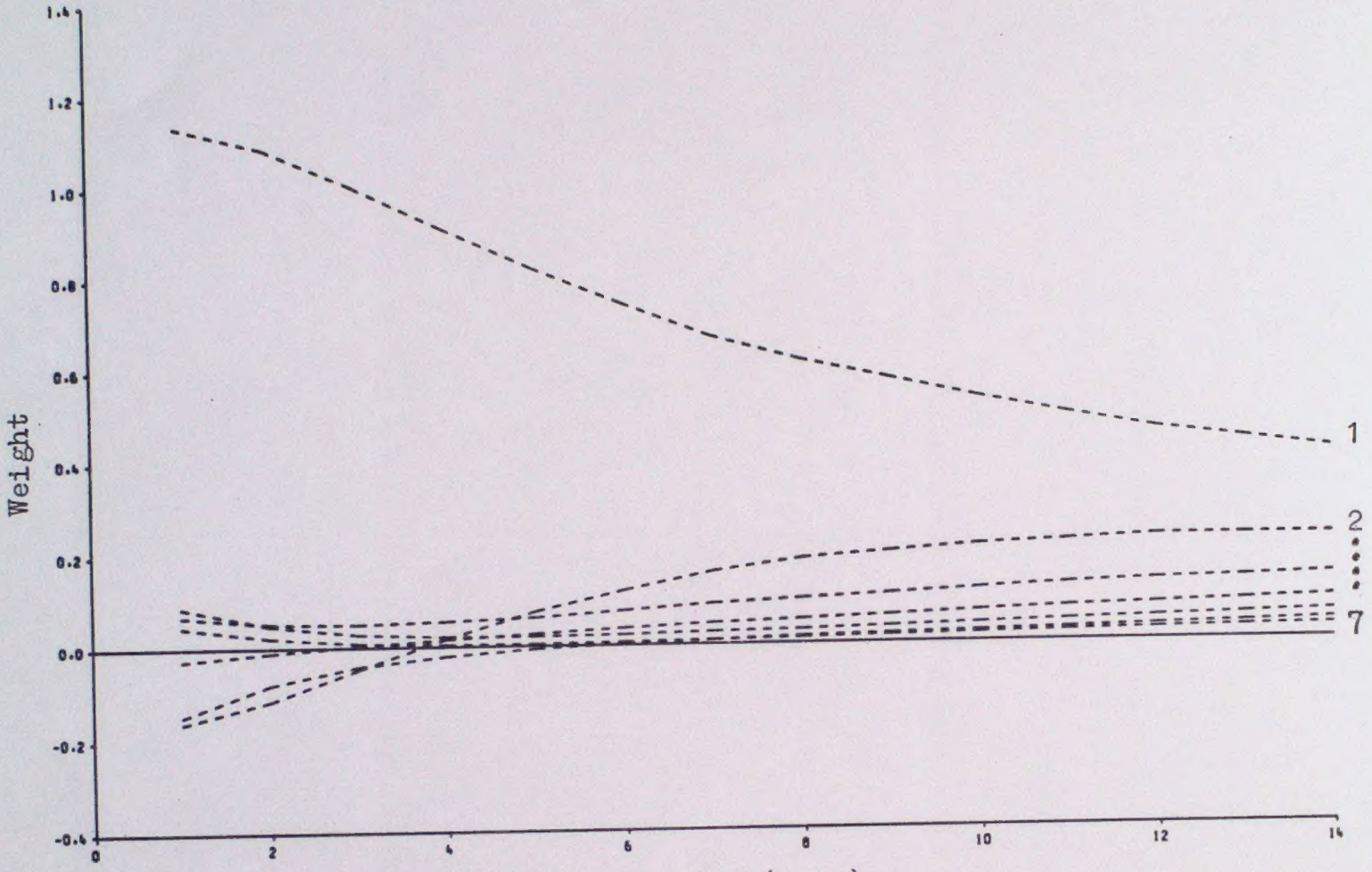
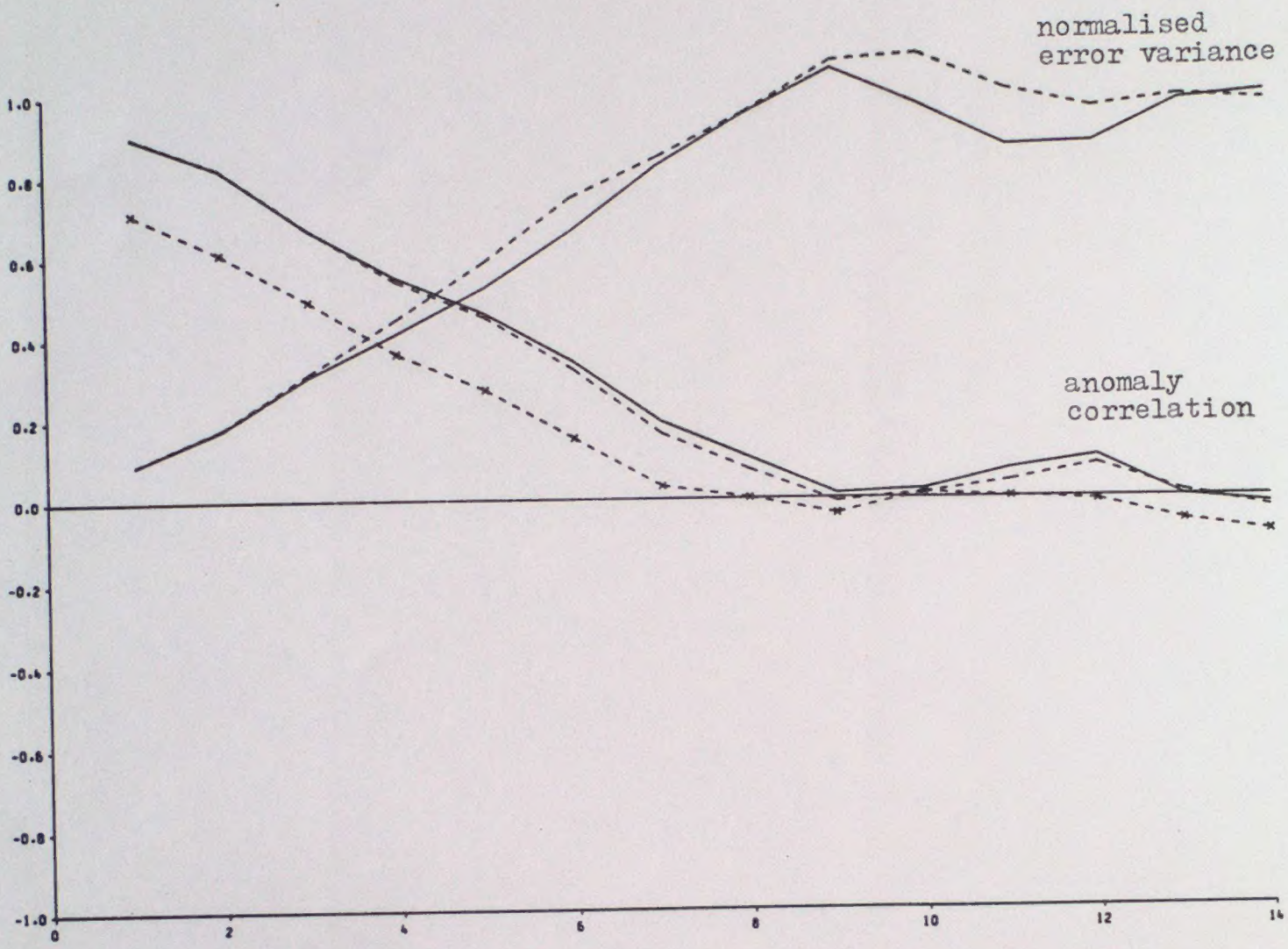

Weight
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
-0.2
-0.4
0
2
4
6
8
10
12
14
1
2
7
Forecast time (days)

Figure 3



Forecast time (days)

Figure 4

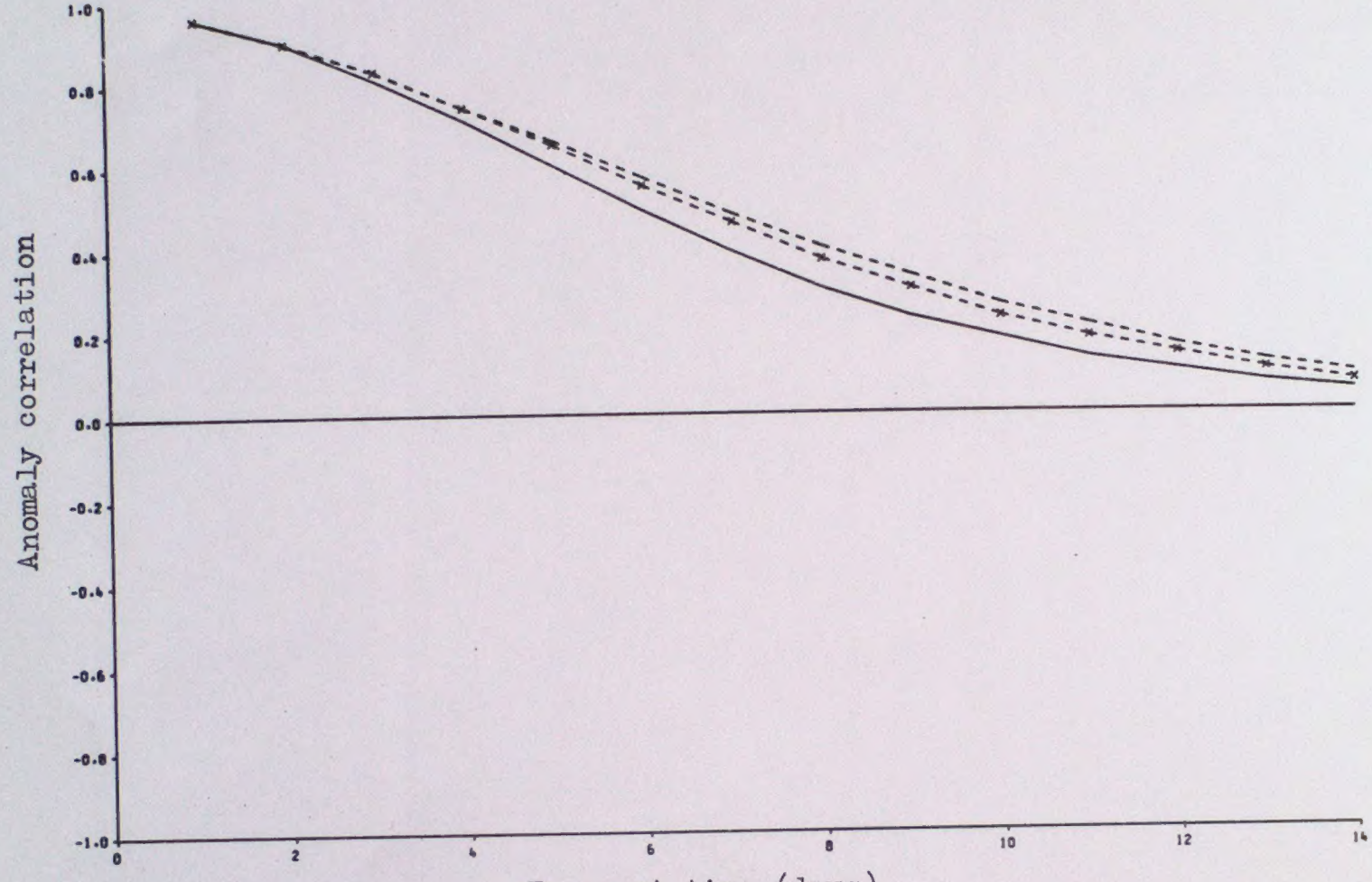

Anomaly correlation
1.0
0.8
0.6
0.4
0.2
0.0
-0.2
-0.4
-0.6
-0.8
-1.0
0
2
4
6
8
10
12
14
Forecast time (days)

Figure 5